DOCUMENT RESUME

ED 105 567 EA 006 983

AUTHOR Churchman, David; And Others
TITLE The Theoretical Basis for Formative Evaluation.
PUB DATE Apr 75
NOTE 14p.; Paper presented at American Educational Research Association Annual Meeting (60th, Washington, D.C., March 31-April 4, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS *Decision Making; *Elementary Education; Evaluation Methods; *Formative Evaluation; Information Utilization; Problem Solving; *Program Evaluation; *Program Improvement

ABSTRACT

One part of the work being done at the Center for the Study of Evaluation has been the development of kits to assist the staffs of elementary schools wishing to conduct evaluations. The third kit in the series provides specific techniques for conducting a formative evaluation, which includes verifying the extent to which a program has been installed as planned and improving the extent to which objectives are achieved. This paper discusses the relationship between the activities called for in the kit and the definition of evaluation as the process of selecting, collecting, analyzing, and reporting information for educational decision-making. Issues raised by evaluation theorists such as problems of objectivity, setting criteria, units of analysis, sampling, unanticipated outcomes, and continuous feedback used to expand this definition as it applies to formative evaluation are also treated in the paper. (Author/MLF)

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED105567

THE THEORETICAL BASIS FOR FORMATIVE EVALUATION

David Churchman
Joseph Petrosko
Laura Spooner-Smith

Center for the Study of Evaluation
University of California, Los Angeles

EA 006 983

Paper presented at the annual meeting of the American Educational Research Association, Washington, D. C., 1975

THE THEORETICAL BASIS FOR FORMATIVE EVALUATION

A framework identifying four major decisions for which evaluation information would be useful was outlined by Alkin in 1967. One part of the work being done at the Center for the Study of Evaluation has been the development of kits to assist the staffs of elementary schools wishing to conduct evaluations for each of the purposes identified in the framework. The third kit in the series provides specific techniques for conducting a formative evaluation, which includes verifying the extent to which a program has been installed as planned and improving the extent to which objectives are achieved.

Introduction

In 1967, Alkin made explicit three assumptions about evaluation and used them to develop both a definition and a framework of evaluation. These assumptions were, first, that evaluation is an information gathering process; second, that the purpose of gathering information is decision making; and third, that evaluation should be presented in a form that the decision maker can understand and use. The first assumption led him to identify the full range of situations requiring evaluation information; the second assumption led him to consideration of the types of information and analyses needed for each situation; and the third assumption led to consideration of the relationship between the evaluator and the decision maker.

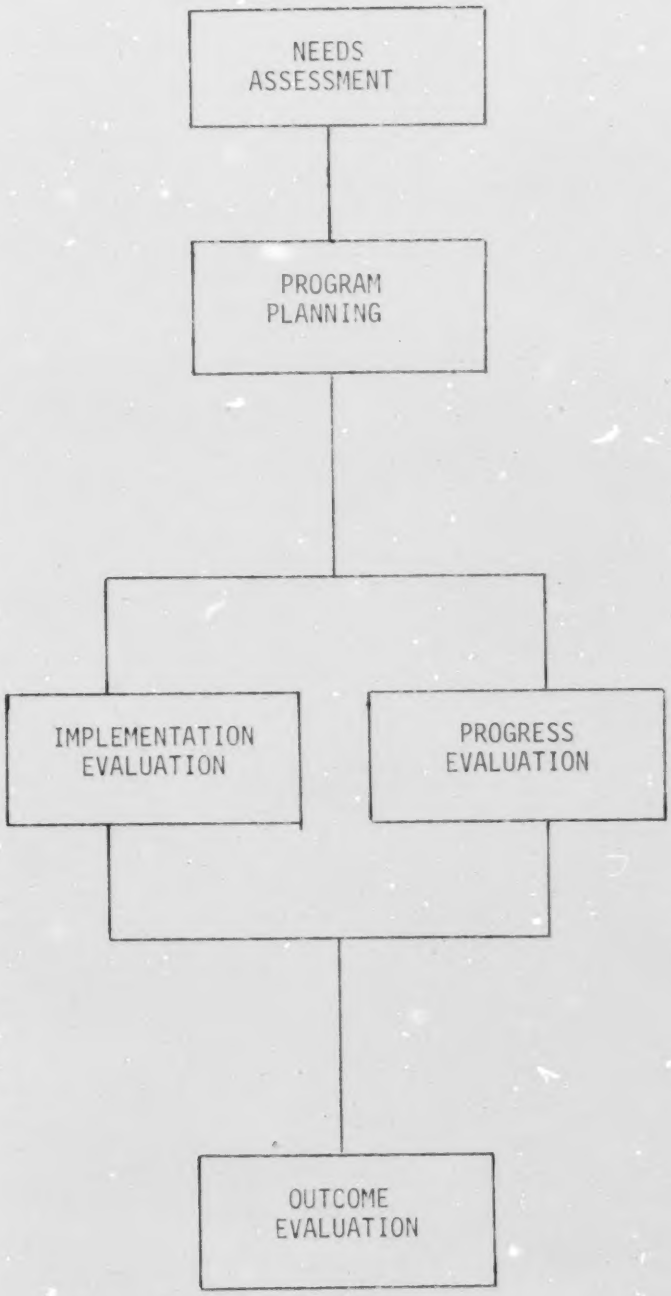
These assumptions and considerations led to both a definition and a framework that provided a conceptual scheme relating evaluation activities, purposes, data collection, data analysis, and reporting to one another. Alkin defined evaluation as the process of ascertaining decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report

summary data useful to decision makers in selecting among alternatives. The framework identified needs assessment, program planning, implementation and progress evaluation, and outcome evaluation as the major decision areas and suggested the major considerations in selecting, collecting and analyzing data for each decision area. Each stage in the framework is shown on Figure 1.

Since then, the Evaluation Technologies Program of the Center for the Study of Evaluation [CSE] has been concerned with elaborating each of these evaluation activities. One strategy has been the development of "evaluation kits." Each kit is intended to provide a particular group of educators with detailed instructions for the completion of one of the evaluation activities identified by the framework. To do so, each kit had to meet at least three requirements. First, each kit had to be consistent with the CSE definition and framework. Second, each kit had to be adapted to the needs, resources, and interests of the user. Third, each kit had to provide information that enabled the user to make a particular type of educational decision. The papers presented during this symposium take up each of these requirements in turn.

Description of the Kit

The CSE Formative Evaluation Kit is intended for use by elementary school principals and their staffs. Assuming the worst, we were determined to design a kit for use by people who had limited time and money for evaluation activities, were conducting an evaluation only because they had to, had no statistical or research skills and no comparison groups available anyway, could get no additional help from outside consultants or district-level staff, and were subject to intense community pressures. Hopefully, most evaluations



NEEDS ASSESSMENT involves stating the objectives to be met and determining how well an existing program is meeting these objectives. This information is used to identify school or program needs.

In PROGRAM PLANNING, the evaluator provides the project director with tools to help make planning decisions. He also builds into the program the procedures that will be needed for assessing whether or not it is operating as planned and how well it is achieving its objectives.

Formative Evaluation includes two elements:

IMPLEMENTATION EVALUATION is a monitoring process to determine the extent to which the specified elements of the program have been implemented as planned.

PROGRESS EVALUATION provides information about the progress of the program's components in meeting the program's objectives. This information is used to modify the program where necessary.

OUTCOME EVALUATION provides information about the success of the entire program. This information can support a decision to maintain, modify, expand, or discontinue the program.

Figure 1

CSE Evaluation Framework

are not conducted under such bleak circumstances. But, a kit designed with these restrictions was more likely to be usable than one assuming a perfect situation.

Developing materials for conducting a formative evaluation consistent with Alkin's framework and appropriate to elementary school users led to a three-volume kit. The first is a guidebook that provides step-by-step instructions on how to conduct an evaluation. The second volume distinguishes materials as a kit rather than a book and is in reality a container for a number of special devices to aid in completing certain steps and a number of forms printed on xeroxable ditto masters on which decisions and evaluation information are recorded. The third volume includes a guide to developing instruments when commercial measures are inappropriate, a glossary of terms used in the kit, and a highly selective annotated bibliography for the individual who wishes to pursue evaluation topics further. Prewrapped for mailing, the kit is designed to sell for approximately thirty dollars.

In essence, the Formative Evaluation Kit breaks the evaluation process down into manageable tasks, provides a logical order for their completion, and includes reusable materials that both reduce overall evaluation costs and provide a complete record of evaluation activities. Several of the steps are related to the administration of an evaluation and deal with such matters as adapting the scope of the evaluation to budget and time constraints. The remainder of the steps stem from theoretical considerations that are the topic of the remainder of this paper.

THE CHARACTERISTICS OF FORMATIVE EVALUATION

The primary purpose of formative evaluation is program improvement.

Formative evaluation is often distinguished from summative evaluation as being concerned primarily with improvement rather than with making final judgments about the worth of a program. Formative evaluation implies intervention and change on the basis of recommendations through which the evaluator is likely to become as committed to the program as other staff members.

This view is in accord with that taken in the CSE Kit with respect both to the purpose of formative evaluation and with the relationship of the evaluator to the program. The position on the latter point is particularly fortunate in view of assumptions about the resources likely to be available for evaluation in elementary schools. If the evaluator is likely to become identified with the program, there is no reason that he cannot be a member of the program staff from the beginning. Rather than requiring selection of an external evaluator, the kit concentrates on distinguishing among roles, permitting a single individual to act in more than one capacity. Instructions in the kit tell the person designated as evaluator how to select, collect, and analyze information so that the staff can reach decisions as to how to improve the program.

Evaluation is decision oriented.

CSE is associated with the decision orientation to evaluation. The purpose of the kit is not to add to the endless arguments as to whether or not this is the "best" approach to evaluation. Rather, the kit provides an opportunity to develop and explore the potential of the decision orientation.

Actual use of the finished kit would contribute to the larger debate by providing data that is usually lacking from the arguments.

Alkin's assumptions, definition and framework were taken as the basis of the kit. The framework identified the purposes for which formative evaluation was to be conducted. The activities necessary to accomplish these purposes were deduced directly from the definition and became the sixteen basic steps into which the kit is divided. This deductive approach ensured that the activities called for by the kit have the potential to provide program directors with the type of information needed to make decisions leading to program improvement, as well as ensuring that the kit is consistent with the model from which it was developed.

The types of evaluation identified in the CSE framework have been deduced from the reference in Alkin's definition to "decision areas of concern." Similarly, the kit identifies three "decision areas of concern" with respect to formative evaluation. These decision areas identify the features of a program that can be modified to improve an educational program. The purpose of collecting evaluation information is to enable the program director to determine which, if any, of these program characteristics require modification. The first of these decision areas, termed preconditions (Stake would have said antecedents), includes the characteristics of all students, staff, materials, equipment, and facilities that constitute the program. The second of these decision areas, termed processes (Stake would have said transactions), is concerned with activities and interactions among students, staff, materials, equipment, and facilities while the program is in operation. The third of these decision areas, termed outcomes (Stake would have said the same), is concerned with both intended and unanticipated effects of the program.

The major purpose in identifying these decision areas is to suggest a broader range of possible program changes than usually is considered. The most common approach to formative evaluation is measurement of en route objectives and repetition or modification of program processes if predetermined criteria are not reached. The approach advocated in the kit permits this interpretation but suggests that other patterns be considered. First, consideration of such changes as improved inservice training for teachers, substitution of different types of equipment, or changes in student entry skills stems from the focus on preconditions. Second, consideration of unanticipated outcomes and the possibility of changing criteria (the logical equivalent of changing entry skills, which draws little objection) is invited through attention to outcomes.

The evaluation context characterizes the evaluation.

The context in which the evaluation is conducted determines to a large extent what can and cannot be accomplished by the evaluator. Among the determinants of context are the resources available, the relationship of the evaluator to the program, and the number and type of decision makers who might make use of the evaluation.

The resources available for conducting an evaluation include time, money, and specialized knowledge. As these factors determine the scope of the evaluation that can be undertaken, two steps in the kit assist the evaluator in adjusting his effort to available resources. In developing the kit, a rather cynical view was taken assuming that the evaluator knew nothing about evaluation, that hard-pressed teachers would be unwilling to give much time to evaluation, and that little money would be made available

either. It is unlikely that an evaluation would be undertaken or could be completed in such circumstances, but this assumption led us to develop a kit usable in minimal circumstances, or whenever one type of resource was available to compensate for a shortage in another type of resource. This view contrasts with the usual approach of presenting elaborate methods requiring advanced research and statistical techniques in the hope that the approximations that will result will yield usable results.

The second dimension of the evaluation context is the relationship between the evaluator and the program being evaluated. The usual distinction is between the "internal" and "external" evaluator. Internal evaluators are more likely to be sensitive to both local factors and local pressures than are external evaluators. In practice, the difference between an internal and external evaluator often is dependent on who is making the distinction, and is of little significance during formative evaluation. First, the formative evaluator is likely to become so involved in the program that he will inevitably become an internal evaluator regardless of his initial stance. Second, the resources available in most elementary schools are such that the evaluator is likely to be drawn from the program staff out of necessity.

A third dimension of the evaluation context is the number and type of decision makers who might make use of the evaluation. Although those involved will vary across schools and time, an elementary school formative evaluation is most likely to interest teachers and curriculum specialists --and possibly will interest parents, the principal, and other school staff. The potential variety of the concerns of these different audiences presents

the evaluator with a difficult task. Rather than attempt to address all possible concerns of all possible audiences in all possible combinations, we recognized that however different these people might be--all are more accustomed to meetings and committees rather than statistics and research reports as a means of reaching decisions. The kit makes suggestions appropriate to this situation. The most important aspect of these suggestions is advice on what tasks the evaluator must do himself and what tasks should involve other individuals. The second aspect of these suggestions is advice on how to run meetings so that tasks are completed efficiently, soundly, and fairly. Thus, for example, the evaluator is responsible for actually collecting data and for preparing it for presentation. However, the identification of questions that guide the evaluation requires the evaluator to solicit a comprehensive list of community concerns about the program, and decisions leading to improvements are reached through a modified Delphi technique moderated by the evaluator.

Formative evaluation should make use of many types of instrument.

Formative evaluation must provide data about program preconditions and processes as well as program outcomes if the information is to be useful in making improvements. As tests measure only outcomes, and no instrument for collecting social science data is suitable for every circumstance, formative evaluation requires a wide range of information collecting methods.

Implementation of this idea in the kit required development of a reasonable number of categories describing the full range of instruments, providing a means for choosing among the types for a given situation, and providing a means for selecting or developing a particular instrument of the

appropriate type for use in data collection. Ultimately, the categories would be arbitrary and a matter of convenience rather than an expression of absolute truth. Too many categories would conflict with the need to provide the elementary school evaluator with a means to choose among them. The need was for a taxonomy that was consistent with the other elements of the kit, that covered as many different situations as possible, and that encouraged the user to depend less on tests than is normally the case. Ultimately, the taxonomy involved eight categories of instruments and four factors that had to be considered in selecting among them. The taxonomy itself and the methods of using it will be discussed in the second paper presented today.

Formative evaluation should be an ongoing process.

Pre- and posttesting is usually considered inadequate for formative purposes. Continuous monitoring permits correcting problems as they occur. Delays in measuring objectives makes ambiguous the cause of any failure to achieve. Finally, continuous measurement tends to increase the aspects of the program that are included in the evaluation, and consequently improves the usefulness of the evaluation itself.

Continuous monitoring of a program is not discussed directly. Rather, it is anticipated to be a natural result of following directions in the kit. This expectation is based on four characteristics of the materials. First, eliciting evaluation questions of interest to a variety of decision makers is likely to lead to concern with critical elements occuring at several point in time. Second, deadlines set for obtaining answers to these questions are expected to occur throughout the length of the program and lead to more-or-

less continual information gathering. Third, the time available for collecting information will probably be evenly distributed throughout the program, and the materials for planning use of time will further lead the evaluator to distribute measurement thoughout the program. Finally, the requirement for monitoring any improvements leads to the generation of new questions and new deadlines while the program is in operation.

Formative evaluation should be an iterative process.

Closely linked to the idea that formative evaluation should be continuous is the idea that it should also be iterative. That is, the question of whether or not improvement resulted from change should be questioned.

To achieve this, the kit discusses factors that must be considered in implementing recommendations for program modification. In effect, these are critical issues that lead to a new set of evaluation questions. To deal with these new questions the evaluator must return to the beginning of the kit and work through it again, selecting means of verifying the implementation of the change and the extent to which it achieved intended results, reporting the results to decision makers, and recycling again with respect to any additional changes.

It is apparent that this process might go on forever. However, the amount of improvement after each iteration can be expected to decline rapidly, while the costs of the evaluation effort will remain relatively fixed. When the program director no longer feels that the cost of the evaluation is warranted in terms of the recommendations resulting, the cycle may be broken. This signals both that the program has reached its full potential and that any further evaluation should be summative.

SUMMARY

The CSE Formative Evaluation Kit is an interpretation and expansion of Alkin's work in evaluation theory development. The activities suggested in the kit, and the order in which they occur, were deduced directly from Alkin's definition rather than by the more usual approach of a review of the evaluation literature. As suggested above, the result is consistent with ideas found in this literature. These ideas are seldom discussed in the kit itself. Rather, the kit is highly specific about activities that should be carried out without explaining the rationale behind them. This approach, as well as the departure from basing evaluation on research design, stems from assumptions made about the interests, needs, and abilities of the intended users. Further differences that stem from these assumptions are explored in the second paper presented during this symposium.